DOCUMENT RESUME

ED 068 505 TM 001 843

AUTHOR	Fitzgibbon, Thomas J.
TITLE	The Use of Standardized Instruments with Urban and Minority-Group Pupils.
INSTITUTION	Harcourt Brace Jovanovich, Inc., New York, N.Y.
PUB DATE	[71]
NOTE	9p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Evaluation Methods; Measurement Instruments; *Minority Group Children; Norm Referenced Tests; *Standardized Tests; Test Bias; *Test Construction; Test Results; *Test Validity; *Urban Youth
IDENTIFIERS	Harcourt Brace Jovanovich Test Department

ABSTRACT

The position of the Harcourt Brace Jovanovich Test Department is set forth for the use of standardized instruments with urban and minority group pupils. Concern is with the effectiveness and usability of the instruments by the organization that published them. Renewed and intensified attention is given to certain aspects of the traditional test-development enterprise to ensure the appropriateness of the instruments in view of concern and controversy about testing in relation to various social issues. Some examples are given of specific ways in which the test development procedures have been modified in relation to the matter of content validity. Other areas of concern include appropriate norms as a national frame of reference and testing in written or dictated English pupils whose native language is not English. It is felt that the responsibility of test publishers includes the consideration of test interpretation and use of test results, and efforts in these areas are being expanded. A short list is included of additional publications concerned with the question of effective communication of standardized test information to the community. (LH)

ED 068505

The Use of Standardized Instruments With Urban and Minority-Group Pupils

Thomas J. Fitzgibbon

TM 001 843

Test Department
Harcourt Brace Jovanovich, Inc.

Test Department
Harcourt Brace Jovanovich, Inc.
The Use of Standardized Instruments With Urban and Minority-Group Pupils

Thomas J. Fitzgibbon

This statement will set forth the position of the HBJ Test Department in this matter and describe the specific steps which have been taken and the concern which will continue to be shown by the Test Department in our effort to improve the effectiveness and usability of our test instruments in the evaluation of urban and minority-group pupils.

There is little question that testing today is in a position of visibility and public concern such as it has not experienced throughout most of its prior history on the American educational and social scene. The issues that have engaged us most seriously on the domestic front for the past several years have been those centering around the emergence of minority groups, the impact of the types of education afforded these groups on their aspirations and advancement in American society, and the social consequences of various types of governmental resources. In all of these contexts, test data have been a focus of concern or a major line of evidence in support of one or another type of governmental action. Because the issues are difficult and complicated ones, beset with political and emotional overtones, there has been a predisposition to criticize the test results reported in support of one or another course of action, and an inclination to question or even to repudiate them when they have seemed to run counter to deeply held attitudes or prejudices. It is difficult to recall a period in which test results have been so surrounded by controversy, or when they have generated such intense reaction, either positive or negative, from large segments of the citizenry. One could cite many instances which bespeak the growing conviction on the part of minority groups that tests constitute for them

an unfair obstacle to advancement in our society. Their responses range from demands for elimination of such testing to demands for development of more appropriate instruments or of modifications in the ways in which measures of performance are interpreted and used in selection decisions.

In the face of this concern and controversy about testing in relation to various social issues, we are conscious of the special responsibilities which fall upon the makers and distributors of tests. As test publishers, we are giving renewed and intensified attention to certain aspects of the traditional test-development enterprise in order to ensure the appropriateness of the final instruments for the uses to which they will be put. We believe it is fair to characterize our response not as "business as usual," but as a more sensitive and more sharply focused conduct of our usual business – that of making good tests and facilitating their proper use.

The above statement of our corporate philosophy in regard to our responsibilities as test publishers in the matter of minority concerns is not new; it was publicly expressed by Dr. Roger T. Lennon, Senior Vice-President of Harcourt Brace Jovanovich, Inc. in a Symposium on *Testing as a Social Problem: Issues and Responsibilities.* American Psychological Association, Washington, D.C., August 31, 1969.

It is the content area which poses the most problems to mutual understanding and acceptance of tests. In most instances, the protest does not concern the objective being measured; rather it is the way the question is asked – the context in which the objective is measured. The context within which the question is asked is part of what we, in the measurement field, call content validity. Most of the complaints about specific questions in current standardized tests fall into two categories: 1) the question will psychologically or emotionally "turn off" a youngster or 2) the context of the question does not relate to the youngster's experiential background. These criticisms are, in many instances, well taken. Until fairly recently, most standardized tests were

constructed by white middle-class people, who sometimes clumsily violate the feelings of the test-taker without even knowing it. In a way, one could say that we have been not so much culture biased as we have been "culture blind."

Some concrete examples of the ways in which we have modified our test development procedures in relation to this matter of content validity are as follows:

1. We attend closely to the set of criteria developed by the Committee on Racial Equality of the Great Cities Research Council for evaluation of instructional materials. These include such matters as sensitivity to prejudice, stereotypes, and material offensive to particular groups; over- or under-emphasis on the worth, capability, or importance of particular racial, religious, or ethnic segments of our national population; inclusion of minority as well as majority examples of worthy Americans in all areas of life and culture; and delineation of life in contemporary urban environments, as well as rural or suburban ones. The assumption is that such attention will enhance the content validity of the tests for minority (and majority) pupils.

2. We have included on our professional staff, both in full-time and in consultant roles, persons who can help us avoid "culture-blindness." Materials in tests under development are reviewed by Blacks, Spanish-speaking Americans, and others who are familiar with the needs and styles of pupils from a variety of minority backgrounds, soliciting their reactions to any content which might be, unintentionally, inappropriate or offensive for such children. (Art work as well as printed material is reviewed in this manner.)

   In the current edition of the Metropolitan, therefore, there is artwork depicting urban scenes, as well as pictures including youngsters and adults who are clearly minority group members. Several items and bits of artwork were revised to eliminate the "mom, dad, two kids, plus dog" stereotype of family composition. In many

stories the adult figure is an uncle, grandfather, or other model, not "Father." In some paragraphs, reviewers found the term "black" used in ways which might be considered derogatory. In such cases, the paragraphs were either scrapped or revised to eliminate the affront. Thus, a story about a crow (described as a big, black bird) who stole shiny dimes was rewritten to eliminate the possible connection of "black" and "theft." These are, of course, only a few instances of the type of reviewing we put our tests through. And we are increasing our efforts along these lines in the reviewing of prospective items for the 1973 edition of the *Stanford Achievement Tests.*

3. We are conducting a number of special tryouts and research projects using test materials with groups of minority pupils, to permit us to observe more directly how well the material functions with them. For example, we are currently undertaking a study in New York City regarding patterns of response to reading and math items in the *Metropolitan Achievement Test*, 1970 Edition. This study is designed to answer the question, "Do children from Title I and non-Title I schools, achieving the same raw scores on a test, obtain that score by answering correctly sets of questions which are significantly different one from the other?" (This research will also concern itself with the question of whether guessing on the test is, in fact, random guessing.)

4. With tests whose results are frequently used to predict future achievement, we have been studying the performance of various subgroups separately to see whether the tests are differentially predictive for the several groups. Included in this category is considerable research concerning the predictive validity of the *Metropolitan Readiness Tests* and the *Murphy-Durrell Reading Readiness Analysis*, both HBJ publications. Much of this data has been published. (See Mitchell,

Blythe C. *Predictive Validity of the Metropolitan Readiness Tests and the Murphy-Durrell Reading Readiness Analysis for White and for Negro Pupils.* Educational and Psychological Measurement, 1967, 27, 1047-1054.) Much additional data is available, as well, from the Test Department, regarding the performance of other minority groups on the *Metropolitan Readiness Tests.*

Another area of test development of great concern to educators is that of appropriate norms. With few exceptions, the norms provided by Harcourt Brace Jovanovich, Inc. for the tests we publish are "national" norms, developed to reflect the performance of the total national population on the instrument being standardized. We can find no consensus as to the utility or helpfulness of separate norms for minority groups on tests such as ours. It seems to us the crucial thing is that all minority groups be properly represented in the national norm populations, and this we are striving to do. The charge that the norms for most tests are based entirely on the performance of middle-class whites simply is not true. In fact, on the standardization of the *Metropolitan Achievement Test*, 1970 Edition, the representation of various minority groups was somewhat higher than their representation in the 1960 population, according to U.S. Census data. When 1970 census data become available, new comparisons will be made and reported.

We feel strongly that "national" norms have meaning and importance for all school systems, regardless of the ways in which a system's pupils may be different from the total population. National norms do describe one reality—the typical performance of the nation's school children. In this respect, they form an important frame of reference; they are a point of departure for decision-making. It should be understood, however, that, by virtue of their being a representative sampling of all the wide variations found within the United States, national norms present a composite picture of national performance, and there are few communities indeed in the entire nation which match that composite picture across all comparisons. (The national average is 2.3

children per family, yet no family has a fractional child.) In many communities, furthermore, the discrepancies between local characteristics and those represented as the typical (by the "norm") are so great as to lead people to conclude the comparison is worthless or perhaps even dangerous.

It is our view that the danger is in using only one frame of reference; one yardstick, when many are available and when each one adds a different dimension to the picture. HBJ does collect and publish information concerning the performance of specific groups, including the large cities, on our various tests, and we encourage users of our tests to interpret test results in terms of several frames of reference, wherever possible. We provide a scoring service designed to enable a school system to interpret their test scores in terms of a local frame of reference. We also provide item analysis service for those schools which desire to use such data for criterion-referenced interpretation of their test results.

A continuing concern with regard to all types of testing has been the question of testing in written or dictated English those pupils whose native language is not English. By far the largest group of non-native English speakers in this country consists of those whose mother tongue is one or another variant of Spanish. Included in this group, however, are Cuban, Puerto Rican, and several dialects of Mexican Spanish. We have been concerned for some time with the question of developing Spanish adaptations (*not* necessarily translations) of a number of our test instruments. However, differences among dialects of Spanish, not to mention cultural differences among Spanish-speaking minorities themselves, have made such projects extremely difficult. Among the tests currently being adapted are the *Otis-Lennon Mental Ability Test* and the *Stanford Early School Achievement Test*. The new *Columbia Mental Maturity Scale* includes in its manual directions in Spanish. This trend will be continued as ways are found to develop appropriate Spanish adaptations of the tests.

As test publishers, we believe our responsibility does not end with the development of appropriate instruments, but extends to the consideration of test interpretation and use of test results. There have been flagrant examples of misuse of test information. We, in turn, are moving toward more extensive and, it is to be hoped, more comprehensible treatment, in our test manuals and accessory materials, of such matters as the nature of tests, proper (and improper) inferences from test results, and cautions and limitations in regard to test interpretation. In particular, we are trying to add more information about how test results can be meaningfully used in a number of concrete situations in the urban centers, in specially funded projects, and the like. We have also greatly expanded our advisory/consultant services to school systems with large numbers of minority group pupils, especially in the large urban centers. We have established, on a pilot basis, a "New York City Desk," with a full-time staff dedicated to helping the city and its local districts plan for and execute testing programs, and evaluate their test results in order to take appropriate action. As we learn from this pilot experience, we shall surely expand our efforts in these areas.

As a test publisher, we sponsor jointly with professional groups across the country Invitational Conferences on Measurement in Education. Many of these conferences in the last few years have been focused directly on the questions of minority concerns with the use of standardized tests. We also publish a number of monographs, articles, and service notebooks concerned with appropriate uses of tests and test results. For instance, one such publication, to be released in 1972, is concerned with the question of effective communication of standardized test information to the community.

Other examples might include the following titles:

"Content, Constructs, Criteria, and Confusion," by Thomas J. Fitzgibbon, Director, Test Department.

Evaluation in the Inner City, Report of an Invitational Conference on Measurement in Education, April 24-25, 1969, Philadelphia, Pennsylvania.

"Reading Tests and the Disadvantaged," an Invited Address given by Thomas J. Fitzgibbon, Director, Test Department at the IRA Committee Conference held at Indiana University, Bloomington, November 18, 1971.

Testimony of Dr. Roger T. Lennon as Expert Witness on Psychological Testing (in the case of Hobson, et al. vs. Hansen, et al., Washington, D.C. schools, 1966).

"Testing and the Culturally Disadvantaged Child," by Roger T. Lennon, Senior Vice-President, HBJ, as one of a series of lectures on Problems in Education of the Culturally Disadvantaged, sponsored by the Boston School Committee, delivered February 26, 1964.

"Testing Issues and Responsibilities as Seen by a Test Publisher," by Roger T. Lennon, Senior Vice-President, HBJ. Presented as part of a Symposium on *Testing as a Social Problem: Issues and Responsibilities.* American Psychological Association, Washington, D.C., August 31, 1969.

It should be clearly understood that what has been said above is not meant to imply that we feel we are "already doing enough"; far from it. But we are trying to do more to be responsive and responsible in the field of educational measurement and evaluation in the Seventies. In this endeavor we welcome the interest and concern of many groups and individuals. The challenges in this area can only be met through the joint efforts of many. We are actively seeking new alliances with parents, community groups, school boards, school administrations, educational organizations, and many others.